http://doi.org/10.35784/iapgos.2804

SELECTED APPLICATIONS OF DEEP NEURAL NETWORKS IN SKIN LESION DIAGNOSTIC

Magdalena Michalska-Ciekańska
Lublin University of Technology, Department of Electronics and Information Technology, Lublin, Poland

Abstract. *The article provides an overview of selected applications of deep neural networks in the diagnosis of skin lesions from human dermatoscopic images, including many dermatological diseases, including very dangerous malignant melanoma. The lesion segmentation process, features selection and classification was described. Application examples of binary and multiclass classification are given. The described algorithms have been widely used in the diagnosis of skin lesions. The effectiveness, specificity, and accuracy of classifiers were compared and analyzed based on available datasets.*

Keywords: dermatoscopic images, neural networks, melanoma, skin lesions

WYBRANE ZASTOSOWANIA GŁĘBOKICH SIECI NEURONOWYCH W DIAGNOZIE ZMIAN SKÓRNYCH

Streszczenie. *Artykuł zawiera przegląd wybranych zastosowań głębokich sieci neuronowych w diagnostyce zmian skórnych z obrazów dermatoskopowych człowieka z uwzględnieniem wielu chorób dermatologicznych, w tym bardzo niebezpiecznej z nich malignant melanoma. Został opisany proces segmentacji zmiany, selekcji cech i klasyfikacji. Uwzględniono przykłady binarnej i wieloklasowej klasyfikacji. Opisane algorytmy znalazły szerokie zastosowanie w diagnostyce zmian skórnych. Porównano i przeanalizowano skuteczność, specyficzność i dokładność klasyfikatorów w oparciu o dostępne zestawy danych.*

Słowa kluczowe: obrazy dermatoskopowe, sztuczne sieci neuronowe, melanoma, zmiany skórne

Introduction

In modern diagnistics of skin diseases, neural networks find a wide range of continguits. They are slowly displacing classical methods [4, 30]. These include methods based on: ABCD, Hunter, Menzies method [25], 7-point checklist [3], TDS, scale Glasgow, scale Hunter and many others [2, 6, 18].

Artificial intelligence is widely used in supporting diagnostic tools in dermatology. Deep learning and machine are effective in distinguishing melanoma from benign skin lesions based on clinical dermatoscopic images. Neural networks are also used in the classification process of dermatoscopic images [13, 11, 36].

An overview of the available methods using the networks has been provided in the [35]. In order for the process to run properly, you need to pay attention to many aspects. One of them is the preliminary preparation of dermatoscopic images. Many available dermatological databases contain images with artifacts: medical markers, applied scale, dark hair, air bubbles, are of poor quality (no contrast in them or are overexposed), the image does not cover the entire outline of the lesion. The next stage is to prepare an appropriate database of training and test images for a given class. Images previously diagnosed correctly by doctors based on a biopsy.

Deep neural networks consist of many layers, each of which identifies more complex elements of the input image. Figure 1 shows an example of a four-layer structure that analyse dermatoscopic images. The input picture of the skin lesion can be classified as a benign birthmark, or melanoma. The layers of the network are interconnected, they have been given appropriate weights. Each node (blue circle) in a layer is connected to each node in the next layer. In fact, the deep network is a much more complex structure than this shown example in figure 1.

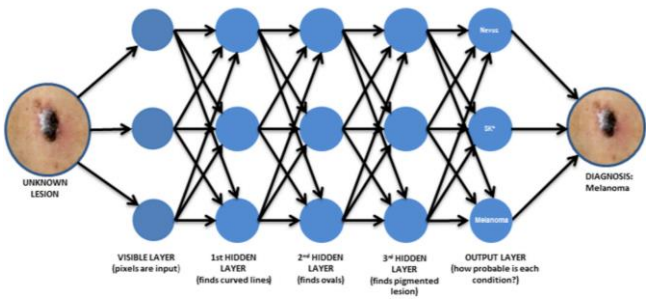

Fig. 1. Diagram of building a neural network [26]

Neural networks contain deep learning algorithms [8], deep convolutional neural networks DCNN [20], synergic deep learning. Great effectiveness achieve algorithms based on neural networks. A more complex structure becomes more effective for data with greater diversity. Also are made hybrid systems [37], witch connecting a model combining synergistic models and deep convolutional neural networks.

1. Segmentation and classification methods

Melanoma occurs in different areas of the body on the neck, torso, arms or legs. The resulting dermatoscopic images are of different quality and contain many artifacts. Therefore, for this, segmenting the entire skin lesion from the image is not an easy task. Traditional segmentation methods such as: adaptive thresholding, Otsu's thresholding, level set active contour, region growing become insufficient.

Increasingly, they are used for this purpose DCNN to train and perform region segmentation in melanoma test images [12, 27, 28]. Images with disease are used to trained model. Training images do not even require often preprocessing. For this purpose, they were used a hybrid deep learning approach based on CNN and recurrent neural network (RNN) [5]. Often researchers use use deep learning [9, 36] and a Deep Residual Network (DRN or ResNet). Efficiency of implementation deep region based convolutional neural network (RCNN) with Fuzzy C-mean (FCM) clustering was checked in [27]. Before the segmentation stage, regions of interest (ROI) are designated, which includes the skin lesion. Containing a fully convolutional neural network (FCN) and a specific convolutional neural network (CNN) in [5] allowed to obtain the result of accuracy 92%, a specificity 93% and a sensitivity 94%. The 28-layer FCN structure made segmentation and with a mask for (ROI).

In [19] researchers implemented neural networks for segmentation, superpixel masks for dermoscopic feature extraction and annotations for images classification procces. Dermatoscopic images from the ISIC database have been segmented and classified using multi-scale fully-convolutional residual networks and a lesion index calculation unit (LICU). Images are classified into three categories: melanoma, seborrheic keratosis and nevus. Scientists have developed framework is named as Lesion Indexing Network (LIN) shown in figure 2. It consists of two FCRN and a calculation unit for lesion index. The next step is bilinear interpolaction for two different images size to get SUM. LICU is result of algorithm, which include possibility and distance maps. The JA and AUC of LIN for segmentation and three categories classification get 0.753 and 0.912.

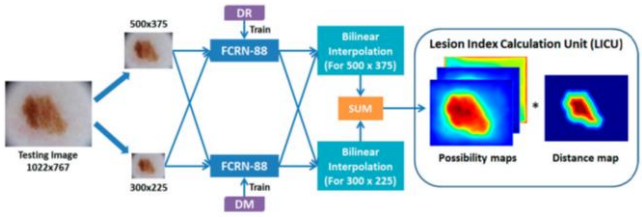

Fig. 2. Lesion Indexing Network (LIN) implementation in [19]

However CNN-based framework, named Lesion Feature Network (LFN) is responsible for dermoscopic feature extraction. The effectiveness of the implemented networks with different number of layers and different parameters, such as batch normalization, weighted softmax, was also studied. In [38] CNN was used to image feature learning and test segmentation results. Figure 3 shows the results of the segmentation algorithm. Binary images are test and real images, designated TP, FN, FN and TN. Classification was made in [10] using CNN based on over 120,000 clinical images.

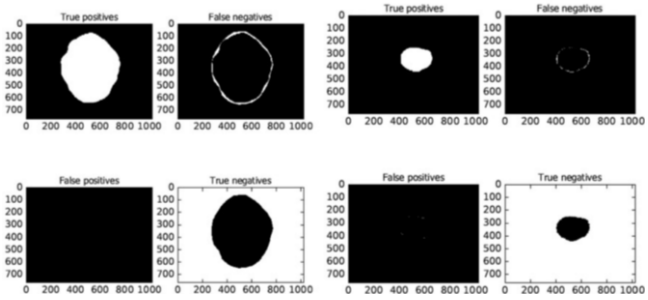

Fig. 3. Results of melanoma test segmentation (FP, TN, TP, FN) [38]

2. Lesions classification process

Deep learning is the basis for numerous applications in dermatology. International projects are also being created to collect a diverse database, and then finding an effective classification method for skin lesions. Such projects are of great interest among scientists. It is made multiclass classification and binary classification. Binary assumes only two classes: melanoma versus benign nevi or benign nonpigmented skin lesions [14]. More extensive classifications allow you to assess more skin diseases causing the formation of various birthmarks.

2.1. Binary classification

Research teams most often work on binary classification based on extensive databases of dermatoscopic images. In most works, the number of training images reaches 900 or even 2000. The images were previously diagnosed by doctors. Their diagnoses were confirmed by histopathological examinations. Gathering such large databases is not easy, it requires a lot of cooperation between scientists and dermatologists.

The mainly diagnosed skin disease that causes the formation of skin lesions is maliganat melanoma. It is diagnosed most often versus no-melanoma lesions or begin nevi. Neural networks are trained on two sets of data, marked as healthy and sick. Table 1 cites six works based on binary classification.

The cited works are characterized by a high level of sensitivity, AUC and specificity. In [29], images from mobile phones were even tested so that anyone with it could make a preliminary diagnosis of lesion using an application on the basis of a photo taken of the property.

Table 1. Binary classification results for dermatoscopic images [35]

Study	Dataset/No. of images	Classification task	Algorithm
[9]	1,279 (900 train, 379 test)	Melanoma versus melanocytic nevi	82% sensitivity, 62% specificity, AUC 0.84
[22]	1,279 (900 train, 379 test)	Melanoma versus melanocytic nevi	82% sensitivity, 76% specificity AUC 0.86
[12]	>100,000	Melanoma versus benign melanocytic nevi	AUC 0.86 (more difficult test-set-100); AUC 0.95 (test-set-300)
[7]	12,378	Melanoma versus atypical nevi	74.1% sensitivity, 86.5% specificity
[17]	Training set: 4,204 biopsyproven melanoma and nevi (1:1) Test set: 804 biopsy-proven melanoma and nevi (1:1)	Melanoma versus nevi	82.3% sensitivity, 77.9% specificity
[29]	Training set: not reported Test set: 551 biopsied lesions (including 125 melanoma) and 999 control lesions (assumed benign)	Melanoma versus nonmelanoma	100% sensitivity, 64.8% specificity with iPhone 6s images

2.2. Multiclass classification

Increasing interest in the processes of diagnosing skin lesions using artificial intelligence has resulted in the development of its use for the diagnosis of many skin diseases. More extensive network models were created, allowing for classification 3, 5, 7 and even 10 different disease. Of course, the introduction of more classes is associated with increased difficulties in solving the algorithm. Multiclass classification mostly helps diagnose nevus, dermatofibroma, melanoma, vascular lesions, benign keratosis, solar lentigo, benign keratosis. Examples of applications for this type of classification are presented in table 2.

Table 2. Multiclass classification results for dermatoscopic images [35]

Study	Dataset/No. of images	Classification task	Algorithm
[23]	Training set: 2,000 Test set: 150	3 disease classes (SK, melanoma, and nevus)	76% sensitivity, 85% specificity AUC 0.87
[24]	Training set: 11,444 Test set: 300 biopsy-verified	5 disease classes (AK, intraepithelial carcinoma, benign keratosis, melanocytic nevi, and melanoma)	AUC 0.96 macro-mean AUC for multiclass AUC 0.93 for benign versus malignant
[33]	Training set: 10,015 Test set: 1,195	7 disease classes (intraepithelial carcinoma including AK and Bowen's disease; BCC; benign keratinocytic lesions including solar lentigo, SK, and LPLK; dermatofibroma; melanoma; melanocytic nevi; and vascular lesions)	81.9% sensitivity, 96.2% specificity (top three algorithms of 139 challenge submissions)
[15]	Images from multiple sources	10 disease classes (nevus, angioma/angiokeratoma, SK, dermatofibroma, solar lentigo, AK, Bowen's disease, melanoma, BCC, and SCC)	95.0% sensitivity, 80.4% specificity for benign versus malignant

In [15], the highest number of class disease was diagnosed. Figure 4 shows boxplots ranked from begin to more malignat categories. Awarded probability scores in scale from 0 (lower probability of malignancy) to 1 (higher probability of malignancy).

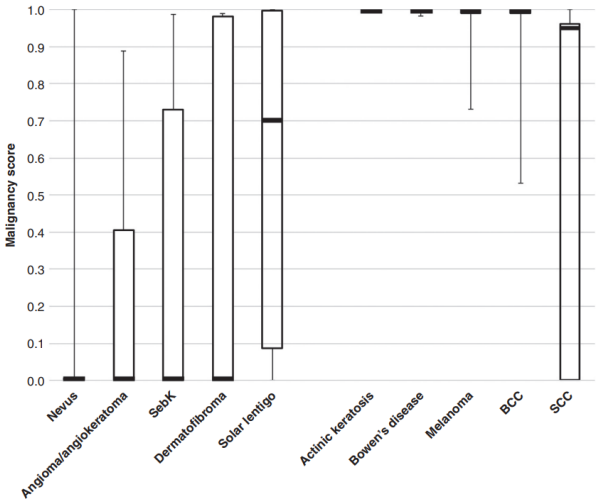

Fig. 4. CNN's melanoma probability scores for benign and malignant categories [15]

3. Features extraction

Convolutional networks have achieved great success in the features selection. There are three types convolutional neural network. First of them are convolution. Mostly they are sets of filters, which are learned for feature extractor to get representative map activation. Next to Pooling – statistically represent the previous layer, that helps to significantly reduce without loss images. The last one is Fully Conected Layer [1].

It is currently being used more and more to smaller receptive window size and smaller stride of the first layer. During training and testing the networks are used the whole image and multiple scales. It is also worth paying attention to the number of layers in the structure of the entire network – its depth. The number of layers should be selected according to the function performed and use of very small convolution filters in all layers. Another important element to fixing other parameters. Networks apply to other sets of images. They achieve excellent performance. That's way convolutional filters numbers of kernel size is kept on maximum small level. Figure 5 shows sparse CNN – Module inception construction. Networks uses convolutions of different sizes to capture details on scales (5×5, 3×3, 1×1).

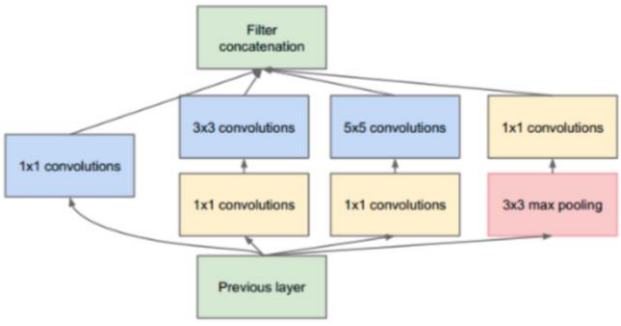

Fig. 5. Module inception construction [32]

In [34] in the first stage, the algorithms were given a modified 7-point method based on the ResNet-50 based supervised deep learning networks (figure 6). Obtained extracted features from dermatoscopic images. Next step is multimodal deep learning framework for automatic malignat melanoma and begin lesions detection by combining deep convolutional neural networks. Based on the obtained features was made clinically constrained classifier chain (CC).

In [34] in the first stage, the algorithms were given a modified 7-point method based on the ResNet-50 based supervised deep learning networks. Reached extracted features from dermatoscopic images. Next step is making multimodal deep learning framework for automatic malignat melanoma and begin lesions detection by combining deep convolutional neural networks. Based on the obtained features got clinically constrained classifier chain (CC).

Commonly known neural network structures are being modified, but new methods are also being used to support them. The applications of new methods are interesting. In [22] a new method called Lesion-classifier. It is based on pixel-wise binary classification on melanoma and non-melanoma cases. All model based on deep convolution encoder-decoder network is shown on figure 7.

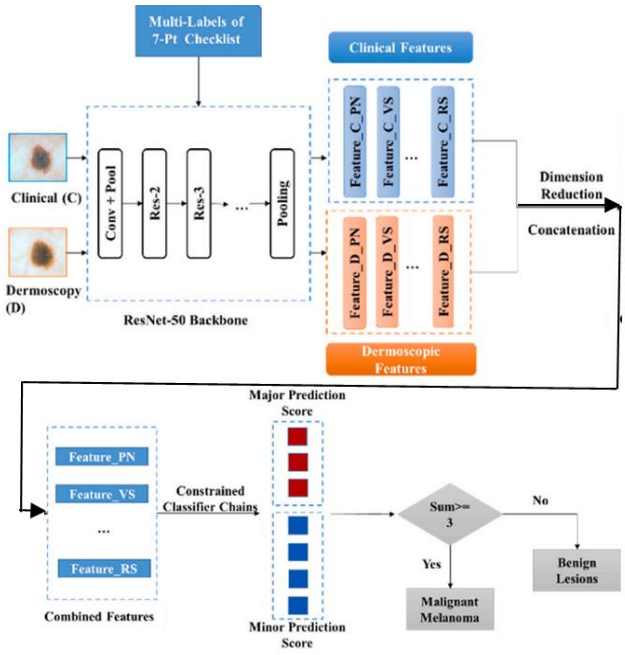

Fig. 6. Operation of the algorithm based on deep learning networks for features extraction [34]

1: **procedure** Encoder(X) ▷ X: x (L, B, H) is an input image with dimension L, B, H
2: Extract the feature map F_map from the input image;
3: For $i = 0$ to M-1:
4: Set $F_{ci} = \text{Conv}(F_{map})$;
5: Set $F_{ri} = \text{Relu}(F_{ci})$;
6: Set $F_{pi} = \text{Pooling}(F_{ri})$;
7: if $i <= M$ then
8: Set $F_{pi+1} = F_{pi}$;
9: else return F_{pi}
10: end if
11: end for
12: **procedure** Decoder(F_{pi}) ▷ F_{pi} is the downsampled feature maps
13: For i =M−1 to 0:
14: Set $F_{di} = \text{Upsamp}(F_{pi})$;
15: Set $F_{ri} = \text{Relu}(F_{di})$;
16: Set $F_{ci} = \text{Conv}(F_{ri})$;
17: if $i <= M$ AND i != 0 then
18: Set $F_{pi-1} = F_{ci}$;
19: else return F_{ci}
20: end if
21: end for
22: Predictedpixels= softmaxclassifier(F_{ci}) ▷ F_{ci} which is the output from the decoder is sent to softmax classifier function for pixel-wise prediction
23: Pi= cluster(Predictedpixels) ▷ The predicted pixels are clustered into segmented output
24: Finalresults= Pi ▷ Final Segmented Output Display Finalresults

Fig. 7. Algorithm based on DCN for binary classification from [1]

Many convolutional neural network from Keras library as VGG16 (512), VGG19 (512), ResNet50 (2048), InceptionV3 (2048) has been used for features extraction in many works [16, 21, 31, 32] – table 3. Artificial neural network structures used The best results were achieved by Logistic Regression using features extracted by the VGG19 with accuracy of 92.5% and F1-score (balance between precision and sensitivity) of 80% (the best result). Precision found by InceptionV3 with Logistic Regression was the best (87.88%).

Table 3. Results deep learning architectures with different classificators [21]

Architecture	Classifier	Accuracy	Precision	Recall	Specificity	F1-Score
InceptionV3	Logistic Regression	0.9250	0.8788	0.7250	0.9750	0.7945
	SVM Linear	0.9200	0.8333	0.7500	0.9625	0.7895
	Naive Bayes	0.7800	0.4750	0.9500	0.7375	0.6333
	SVM RBF	0.8700	0.8889	0.4000	0.9875	0.5517
	AdaBoost	0.7800	0.4500	0.4500	0.8625	0.4500
	Random Forest	0.8250	0.6923	0.2250	0.9750	0.3396
ResNet50	SVM Linear	0.9100	0.8438	0.6750	0.9688	0.7500
	Logistic Regression	0.9000	0.7941	0.6750	0.9563	0.7297
	Random Forest	0.8850	0.7931	0.5750	0.9625	0.6667
	Naive Bayes	0.8100	0.5152	0.8500	0.8000	0.6415
	SVM RBF	0.8850	0.9048	0.4750	0.9875	0.6230
	AdaBoost	0.8300	0.5750	0.5750	0.8938	0.5750
VGG16	Logistic Regression	0.9150	0.7949	0.7750	0.9500	0.7848
	SVM Linear	0.9100	0.7750	0.7750	0.9438	0.7750
	SVM RBF	0.8950	0.7317	0.7500	0.9313	0.7407
	Random Forest	0.8950	0.7714	0.6750	0.9500	0.7200
	Naive Bayes	0.8400	0.5667	0.8500	0.8375	0.6800
	AdaBoost	0.8700	0.6944	0.6250	0.9313	0.6579
VGG19	**Logistic Regression**	**0.9250**	**0.8571**	**0.7500**	**0.9688**	**0.8000**
	SVM Linear	0.9050	0.8182	0.6750	0.9625	0.7397
	Random Forest	0.9100	0.8929	0.6250	0.9813	0.7353
	SVM RBF	0.8950	0.8276	0.6000	0.9688	0.6957
	Naive Bayes	0.8150	0.5246	0.8000	0.8188	0.6337
	AdaBoost	0.8450	0.6452	0.5000	0.9313	0.5634

4. Discussion and conclusions

The process of diagnosing skin lesions based on neural networks is an issue of interest to many research teams, many works have been created in this field and further research is still being conducted. Currently, an important problem in the use of deep neural networks is the preparation of an appropriate database of dermatoscopic images, their initial preparation, taking an effective segmentation method and using a well-chosen structure for the lesion classification. In the future, most likely, the classic methods of diagnosing skin lesions based on geometry, shape of the lesion using segmentation will begin to be replaced by diagnostics based on the analysis of lesion texture.

References

[1] Adegun A., Viriri S.: Deep learning-based system for automatic melanoma detection. IEEE Access 8/2020, 7160–7172.
[2] Alendar F. et al.: Clear definitions, simple terminology, no metaphoric terms. Expert Rev Dermatol 3/2008, 27–29.
[3] Argenziano G. et al.: Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. Comparison of the ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis. Archives of Dermatology 134/1998, 1563–1570.
[4] Argenziano G. et al.: Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet. Journal of American Academy of Dermatology 48(5)/2003, 679–693.
[5] Attia M. et al.: Skin melanoma segmentation using recurrent and convolutional neural networks. Biomedical Imaging (ISBI 2017) – 14th International Symposium, IEEE, 2017, 292–296.
[6] Blum H., Ellwanger U.: Digital image analysis for diagnosis of cutaneous melanoma. Development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions. British Journal of Dermatology 151(5)/2004, 1029–1038.
[7] Brinker T. J. et al.: Deep learning outperformed 136 of 157 dermatologists in a head-to-head der moscopic melanoma image classification task. Eur J Cancer 113/2019, 47–54.
[8] Codella N., Cai J., Abedini M., Garnavi R., Halpern A., Smith J. R.: Deep Learning, Sparse Coding, and SVM for Melanoma Recognition in Dermoscopy Images. In: Zhou L., Wang L., Wang Q., Shi Y. (eds) Machine Learning in Medical Imaging. MLMI 2015. Lecture Notes in Computer Science 9352. Springer, Cham 2015 [http://doi.org/10.1007/978-3-319-24888-2_15].
[9] Codella N. et al.: Deep learning ensembles for melanoma recognition in dermoscopy images. 2016 [http://arxiv.org/abs/1610.04662] (accessed 4 October 2019).
[10] Esteva A. et al.: Dermatologist-level classification of skin cancer with deep neural networks. Nature 542/2017, 115–118.
[11] Esteva A.: Dermatologist-level classification of skin cancer with deep neural networks. Nat. Res. 542(7639), 2017, 115–118.
[12] Ge Y. et al.: Melanoma segmentation and classification in clinical images using deep learning. Proceedings of the 2018 10th International Conference on Machine Learning and Computing, 2018, 252–256.
[13] Ge Z. et al.: Skin disease recognition using deep saliency features and multimodal learning of dermoscopy and clinical images. Springer, Cham LNCS 10435, 2017, 250–258.
[14] Haenssle H. A. et al.: Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol 29/2018, 1836–1342.
[15] Haenssle H. A. et al.: Man against machine reloaded: performance of a market-approved convolutional neural network in classifying a broad spectrum of skin lesions in comparison with 96 dermatologists working under less artificial conditions. Ann Oncol 31/2020, 137–143.
[16] He K., Zhang X, Ren, S., Sun J.: Deep residual learning for image recognition, Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, 770–778.
[17] Hekler A. et al.: Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. Eur J Cancer 118, 2019, 91–96.
[18] Kittler H. et al.: Dermatoscopy of unpigmented lesions of the skin: A new classification of vessel morphology based on pattern analysis. Dermatopathology: Practical & Conceptual 14(4), 2008, 3.
[19] Li Y., Shen L.: Skin lesion analysis towards melanoma detection using deep learning network. Sensors 18, 2018, 556.
[20] Lopez A. R. et al.: Skin lesion classification from dermatoscopic images using deep learning techniques. Biomedical Engineering 852, 2017, 852-053 [http://doi.org/10.2316/P.2017.852-053].
[21] Maia L. et al.: Evaluation of melanoma diagnosis using deep features. 25th International Conference on Systems, Signals and Image Processing (IWSSIP), 2018, 1–4.
[22] Marchetti M. A. et al.: Results of the 2016 international skin imaging collaboration international symposium on biomedical imaging challenge: comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images. J Am Acad Dermatol 78, 2018, 270–277.
[23] Marchetti M. A. et al: Computer algorithms show potential for improving dermatologists' accuracy to diagnose cutaneous melanoma: results of the international skin imaging collaboration 2017. J Am Acad Dermatol 82, 2020, 622–627.
[24] Maron R. C. et al.: Systematic outperformance of 112 dermatologists in multiclass skin cancer image classification by convolutional neural networks. Eur J Cancer 119, 2019, 57–65.
[25] Menzies S. et al.: Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features. Archives of Dermatology 132, 1996, 1178–1182.
[26] Murphree D. H. et al.: Deep learning for dermatologists: Part I. J Am Acad Dermatol 2020, 1–9, [http://doi.org/10.1016/j.jaad.2020.05.05].
[27] Nida N. et al.: Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering. International Journal of Medical Informatics 124, 2019, 37–48.
[28] Nijeweme-d'Hollosy W. et al.: Evaluation of three machine learning models for self-referral decision support on low back pain in primary care. Int. J. Med. Inform. 110, 2018, 31–41.
[29] Phillips M. et al.: Assessment of accuracy of an artificial intelligence algorithm to detect melanoma in images of skin lesions. JAMA Netw Open 2, 2019, 1913436.
[30] Rosendahl C., Cameron A., McColl I., Wilkinson I.: Dermatoscopy in routine practice. Chaos and Clues. Australian Family Physician 41(7), 2012.
[31] Simonyan K., Zisserman A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.
[32] Szegedy C. et al.: Going deeper with convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, 1–9.
[33] Tschandl P et al.: Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, webbased, international, diagnostic study. Lancet Oncol 20, 2019, 938–947.
[34] Wang Y. et al.: Incorporating clinical knowledge with constrained classifier chain into a multimodal deep network for melanoma detection. Computers in Biology and Medicine 137, 2021, 104812.
[35] Young A. T. et al.: Artificial intelligence in dermatology: A Primer. Journal of Investigative Dermatology 140, 2020, 1504–1512.
[36] Yu L. et al.: Automated melano-ma recognition in dermoscopy images via very deep residual networks. IEEE transactions on medical imaging 36(4), 2017, 994–1004.
[37] Zhang J. et al.: Skin lesion classification in dermoscopy images using synergic deep learning. Springer Nature Switzerland, LNCS 11071, 2018, 12–20.
[38] Zhang X.: Melanoma segmentation based on deep learning. Computer Assisted Surgery 22, 2017, 267–277.

M.Sc. Magdalena Michalska-Ciekańska
e-mail: magdalena.michalska@pollub.edu.pl

Ph.D. student at Department of Electronics and Information Technology, Lublin University of Technology. Recent graduate Warsaw University of Technology The Faculty Electronics and Information Technology. Her research interests include medical image processing, optoelectronics, spectrophotometry.

http://orcid.org/0000-0002-0874-3285

otrzymano/received: 29.10.2021 *przyjęto do druku/accepted: 15.12.2021*